

Pattern recognition with neuronal avalanche dynamics

L. Michiels van Kessenich* and D. Berger

Computational Physics for Engineering Materials, IfB, ETH Zürich, 8093 Zürich, Switzerland

L. de Arcangelis

*Department of Engineering, University of Campania “Luigi Vanvitelli,” I-81031 Aversa (CE), Italy
and INFN Sezione Naples, Gruppo Collegato Salerno, Salerno, Italy*

H. J. Herrmann

*Computational Physics for Engineering Materials, IfB, ETH Zürich, 8093 Zürich, Switzerland
and Departamento de Física, Universidade Federal do Ceará, 60451-970 Fortaleza, Ceará, Brazil*

(Received 23 August 2018; published 22 January 2019)

Pattern recognition is a fundamental neuronal process which enables a cortical system to interpret visual stimuli. How the brain learns to recognize patterns is, however, an unsolved problem. The frequently employed method of back propagation excels at this task but has been found to be unbiological in many aspects. In this Rapid Communication we achieve pattern recognition tasks in a biologically, fully consistent framework. We consider a neuronal network exhibiting avalanche dynamics, as observed experimentally, and implement negative feedback signals. These are chemical signals, such as dopamine, which mediate synaptic plasticity and sculpt the network to achieve certain tasks. The system is able to distinguish horizontal and vertical lines with high accuracy, as well as to perform well at the more complicated task of handwritten digit recognition. Resulting from the learning mechanism, spatially separate activity regions emerge, as observed in the primary visual cortex using functional magnetic resonance imaging techniques. The results therefore suggest that negative feedback signals offer an explanation for the emergence of distinct activity areas in the visual cortex.

DOI: [10.1103/PhysRevE.99.010302](https://doi.org/10.1103/PhysRevE.99.010302)

Learning in neural networks occurs through the modification of the synaptic structure of the neuronal system [1,2] and many different mechanisms have been explored in the attempt to explain it. One of the most successful methods is the back-propagation algorithm [3–5] which excels at pattern recognition tasks [4]. It has, however, been stressed that back propagation does not occur in real biological systems [6]. A prominent issue is the *weight transport* problem [7]: Each neuron in the back-propagation algorithm requires the knowledge of the full downstream path through the network to precisely calculate the necessary synaptic changes to minimize the error [8]. Other nonbiological ingredients in back propagation include the fact that the computation of error gradients would need to be precisely clocked between forward- and backward-propagation phases [9]. How biological neural networks modify the synaptic efficacies to achieve learning can therefore not be explained by this algorithm [6]. Back propagation has been modified to avoid some of the unbiological aspects [9]. Lillicrap *et al.* used random synaptic feedback weights which circumvents the *weight transport* problem [8] and Brand *et al.* developed a mechanism to modify synaptic strengths without using a feedback network [10]. Other approaches use, for example, the output of a biologically inspired model to train a linear classifier using

back propagation [11]. Another promising approach involves liquid state machines (LSMs) but some of the aspects of LSMs are difficult to motivate biologically [12–15]. How the brain achieves pattern recognition remains, therefore, an open problem.

Recent experimental and numerical results have suggested that spontaneous brain activity consists of neuronal avalanches with critical features [16–27]. These are bursts of firing neurons whose size and duration distribution follow a power law, indicating that avalanches do not exhibit a characteristic size. This is the fingerprint of self-organized criticality (SOC) [28] and models inspired in SOC have accurately reproduced the experimentally observed behavior of spontaneous brain activity [29–34]. Furthermore, these models are able to learn simple boolean rules through negative feedback signals [35–37]. In contrast to the back-propagation method, negative feedback signals provide a biologically plausible mechanism for learning [38] which is based on the release of specific chemicals, as, for example, a monoamine such as dopamine, which has been found to mediate synaptic plasticity [39–41]. These chemicals are released by specific neurons and diffuse into extracellular space [42]. In this Rapid Communication we explore the capabilities of negative feedback signals to tackle pattern recognition in a neuronal model which exhibits avalanche dynamics and incorporates the main neurobiological mechanisms [43] such as inhibitory neurons [44], synaptic plasticity [45], synaptic fatigue

*laurens@ethz.ch

[46–50], and the neuronal refractory time [51]. The system identifies patterns by exhibiting increased neuronal activity at a specific output region. In case the answer is wrong, chemical signals are released which mediate plasticity and modify the synaptic structure. Over time the plasticity mechanism sculpts the network until it performs the tasks correctly. As soon as the system provides the right response, we investigate the cortical dynamics arising from the input of the specific patterns. From experiments it is known that the primary visual cortex V1 contains specific areas, called orientational columns [52–54]. These are organized regions of neurons that respond to visual line stimuli of different orientations. These orientational columns are subject to synaptic plasticity as shown in visual deprivation experiments [55]. Our results show how such pattern-specific activity regions emerge from the local release of the chemical negative feedback signals, thus providing a possible explanation for the formation of these columns.

Consider a network of N neurons, each characterized by a membrane potential v_i . Neurons are connected via synapses with weight w_{ij} and a fraction p_{in} of them are inhibitory. As soon as at time t the potential at some neuron reaches the threshold v_c , the neuron fires, triggering the release of a fraction u of the available neurotransmitter. It has been shown that the readily releasable pool of neurotransmitter is about 5% of the total available transmitter [49,50], and therefore we set $u = 0.05$. The potential of the postsynaptic neurons j then changes proportionally to the amount of neurotransmitter released according to

$$v_j(t+1) = v_j(t) \pm v_i u w_{ij}(t), \quad (1)$$

where the $+$ and $-$ stand for excitatory and inhibitory presynaptic neurons, respectively. After firing, the synaptic weight w_{ij} will have less neurotransmitter available according to

$$w_{ij}(t+1) = w_{ij}(t)(1-u), \quad (2)$$

and the neuron enters a refractory state for one time step, during which it will not generate further action potentials. If the change in potential in the postsynaptic neurons is sufficient to surpass the threshold, an avalanche of neuronal activity propagates through the network. It stops as soon as all neurons are below threshold and an external stimulation is needed to trigger further activity. During an avalanche the synapses lose some neurotransmitter and a recovery is needed to sustain network activity. In Ref. [43] the synapses recover continuously as randomly stimulated avalanches propagate throughout the network. In the present work the stimulation always occurs at the input neurons which would quickly deplete their outgoing synapses of the neurotransmitter. One option to recover these synapses is to simulate many random avalanches between two input patterns to allow the system's recovery. This is very demanding from the computational point of view, therefore we consider the time between two external stimuli sufficient for the synapses to recover to the original value W_{ij} . In this way the synaptic strength is represented by two variables, the long-term synaptic strength W_{ij} and the short-term synaptic strength w_{ij} .

The network (Fig. 1) consists of an input layer arranged on a grid, an intermediate network, and output regions. In pattern recognition different patterns need to be classified into classes where many similar patterns belong to each class. For each

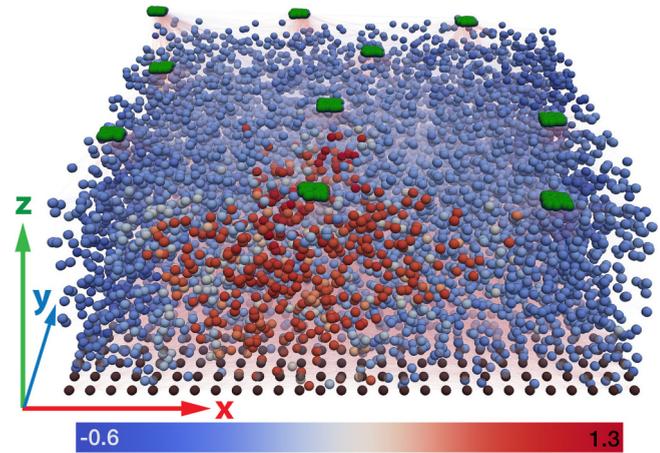


FIG. 1. A network of 8000 neurons with ten output regions (green) and the input grid (black). The intermediate scale-free network has a color map indicating $\langle r_i \rangle_C - \langle r_i \rangle$, the localized activity for an input class C in the MNIST task. $\langle r_i \rangle_C$ is the average activity for the class $C = 3$ and $\langle r_i \rangle$ is the overall average activity of the network. The parameters used are $p_{\text{in}} = 0.3$, $N_O = 50$, $k_{\text{min}} = 10$.

class an output region of N_O neurons is placed randomly, in such a way that they do not overlap, on the side of the network opposite to the input grid. The neurons in the intermediate network are randomly placed in a cuboid whose dimensions are fixed to z_0 in the z direction and the x and y directions scale depending on the number of neurons N in the network to keep the density constant. Maintaining a constant height z_0 for various network sizes N ensures that the chemical distance between input and output does not increase with N , affecting the performance, as has been found in the learning of boolean rules [35]. The numerical value for z_0 is not important as all other distance-dependent quantities will be defined relative to z_0 . In order to choose the network structure, we notice that functional magnetic resonance imaging (fMRI) measurements have reported that the functional network has scale-free properties. Namely, the distribution of outgoing degrees k follows a power law $P(k) \propto k^{-2}$ [56]. The intermediate network is therefore scale free, i.e., the neurons have an outdegree distribution following a power-law decay in the range $k_{\text{out}} \in [k_{\text{min}}, k_{\text{max}}]$. We will vary k_{min} and set $k_{\text{max}} = 100$. The visual cortex has mostly feed-forward connections, i.e., it does not contain cycles [57,58], and we therefore use forward-directed synapses for the entire system, which are connected according to a distance-dependent probability $P(r) \propto e^{-r/r_0}$ with $r_0 = 0.3z_0$ [37].

Neuronal bursts of activity are triggered by stimulating the input neurons according to various patterns. This activity percolates towards the output regions and once the activity ceases, the system's response is evaluated according to the action potentials which occurred in the output regions. If the most active output region, i.e., the output region with the most firing events during the neuronal propagation, corresponds to the correct class, we consider the response as correct and no modification of the synaptic structure occurs. Conversely, if the system made a mistake, all output neurons release the “learning signal.” The output neurons which should have remained silent, but which incorrectly fired, release a negative

feedback signal, weakening the active synapses in the network, reducing their future activity. The output neurons which should have been active but did not fire release a positive signal which strengthens the active connections to increase activity. The signal diffuses through the extracellular space and reaches the synapses in the network. Those synapses which were active during the avalanche will modify according to the error signals received,

$$W_{ij} \rightarrow W_{ij} \pm \sum_n \alpha e_n f(d_{jn}), \quad (3)$$

where α is a parameter which determines the strength of plastic adaptation. Inhibitory and excitatory synapses due to their inverse effect on postsynaptic neurons need to modify their weights differently, according to homeostatic plasticity [59], + for excitatory and - for inhibitory synapses. The value for e_n is either $\{-1, 0, 1\}$, depending on whether the neuron incorrectly fired (-1), behaved correctly (0), or incorrectly remained silent (1). The function $f(d_{jn})$ describes the dependence of the learning signal strength on the distance d_{jn} between the synaptic button j and the output neuron n releasing the error signal e_n . For the function $f(d_{jn})$ we use $f(d_{jn}) = e^{-d_{jn}/d_0}$ with $d_0 = 0.3z_0$. The dependence of the learning performance on d_0 has been studied in Ref. [37], where it is shown that the learning performance is optimized when the learning length d_0 is of the order of the system size. According to Eq. (3), the negative feedback learning mechanism modifies the synaptic strengths solely on the basis of local information, the neuronal activity, and the chemical error signals. In contrast to back propagation, the neurons therefore do not require any global information about downstream paths, nor their position within the network. This rule implies that the concentration of messenger molecules implementing negative feedback, and therefore the strength of the synaptic modification, decays exponentially over distance, i.e., synapses close to the output neurons are affected more strongly than distant synapses. We also explored other types of decay functions, such as power laws or Gaussian functions, confirming that the analytic form of $f(d_{jn})$ does not influence the results significantly.

When a pattern is presented to the network for the first time, it usually generates incorrect classifications and feedback signals are released by the output neurons. These adapt the synaptic strengths and over time the learning mechanism sculpts the structure of the network. We define the learning performance as $P = n_{\text{correct}}/n_{\text{total}}$, where n_{total} is the total number of input patterns presented to the network and n_{correct} is the number of times the system gives the correct response. To evaluate this performance all patterns of the given task are presented to the network in a random sequence. We first investigate how the network learns simple patterns such as horizontal and vertical lines. The patterns used are of size 20×20 with lines of width 3 as shown in Fig. 2. In this case the inputs need to be classified into two classes. To ensure that the task is nonlinearly separable, we add random shifts to these patterns [5], which leads to a total of six different shifted patterns for each class. The task can be made more challenging by adding additional classes such as zigzag lines or a checkerboard. Figure 2 shows how the performance of the network changes when more classes are added to the

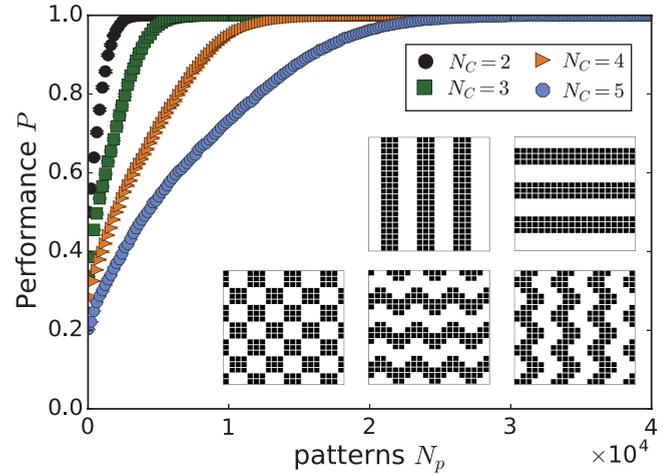


FIG. 2. Performance vs the number of patterns presented to the network. The used pattern classes are shown in the bottom right corner. First, only the horizontal and vertical lines are presented ($N_C = 2$). Other classes are then added to the task to increase the complexity. The parameters used are $N = 8000$, $p_{\text{in}} = 0.3$, $N_O = 50$, $k_{\text{min}} = 10$.

classification task. The number of trials the networks needs to classify the patterns grows exponentially with the number of classes N_C [60]. Our patterns have significant overlap with each other, yet the system can learn to distinguish them with 100% accuracy. We therefore consider next more complicated tasks such as handwritten digit recognition using the Modified National Institute of Standards and Technology (MNIST) database [61] which contains 60 000 images of handwritten digits for training the network and 10 000 images for testing the performance. The network needs to classify the input patterns into ten different classes and therefore ten output regions are needed. Figure 3 shows how the performance changes as a function of the number of training patterns. The performance initially increases rapidly and saturates at around 85%. The maximum performance obtained depends on various parameters which define the network structure [60] and a significant dependency is found on the fraction of inhibitory neurons p_{in} . At low fractions of inhibitory neurons $p_{\text{in}} = 0$ the performance is poor and optimal performance is found in the range of $p_{\text{in}} = 20\% - 40\%$, close to the value found in mammalian brains, confirming previous results on boolean multitask learning [36]. The presented learning mechanism is not intended to compete with back-propagation implementations which with various techniques, such as data preprocessing techniques, different network topologies (convolutional networks), elastic distortions, and more, achieve performances greater than 99% [62–64]. Rather, it provides a learning mechanism which is founded on biological considerations applied to a biological neural network model.

Experimental studies agree that the distribution of neuronal avalanches occurring during spontaneous activity in the resting state exhibits universal critical properties. The avalanche size s can be obtained during pattern recognition by measuring the number of active neurons, as a response to an input. For this we present the 10 000 digits of the test set ($N_C = 10$) to the system and measure the avalanche size distribution $P(s)$

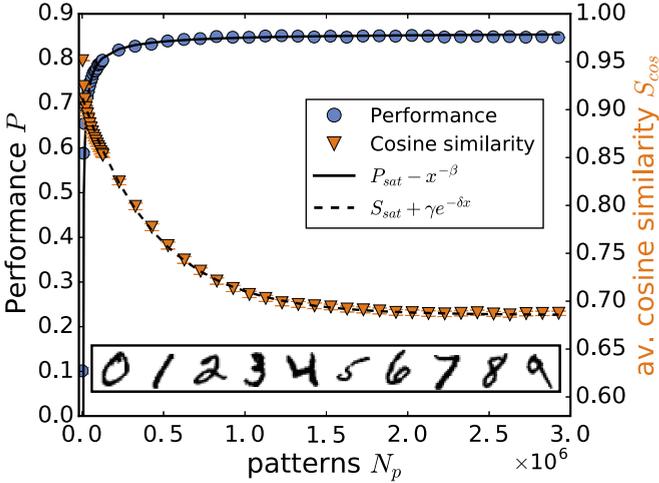


FIG. 3. Performance on the MNIST data set vs the number of handwritten digits (patterns) presented to the network. As the performance increases, the activity vectors $\langle r_i \rangle_C$ become increasingly different from each other, leading to the decrease in the average cosine similarity S_{cos} . The parameters used are $N = 8000$, $p_{\text{in}} = 0.3$, $N_O = 50$, $k_{\text{min}} = 10$, $N_C = 10$. The values presented are an average over 100 configurations. The performance approaches the saturation level $P_{\text{sat}} \approx 0.85$ as $P_{\text{sat}} - N_p^{-0.72}$. The cosine similarity decays exponentially to the saturation level $S_{\text{sat}} \approx 0.68$ as $S_{\text{sat}} + \gamma e^{-\delta N_p}$ with $\gamma = 0.23$ and $\delta = 5.5 \times 10^{-3}$. Some example patterns from the MNIST database are shown below.

shown in Fig. 4. It should be noted that the response to a pattern generates stimulated activity, not spontaneous activity, and therefore no power-law distribution is expected. Rather, $P(s)$ is normally distributed and changes during the evolution of the synaptic weights. At the beginning, when the system gives mostly incorrect responses, the distribution is broader and has its maximum at larger avalanches. An avalanche of size s represents a configuration where s neurons are active and $N - s$ are inactive. In the framework of spin models, this avalanche might be described as a configuration with s up spins and $N - s$ down spins. The probability $P(s)$ to obtain an avalanche of size s , regardless of the individual firing neurons, might be identified as the probability to observe such a configuration. We can therefore define a Shannon entropy as $H = -\sum_s P(s) \log[P(s)]$ and monitor the evolution of

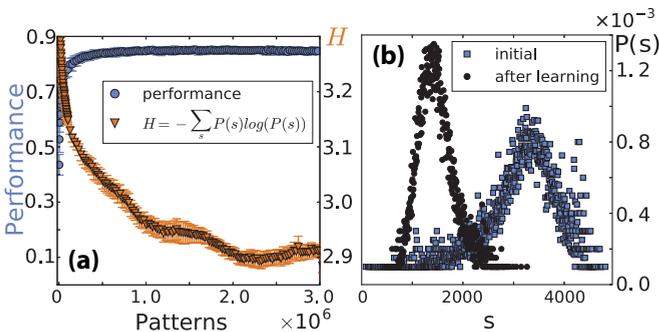


FIG. 4. (a) As the performance increases, the entropy $H = -\sum_s P(s) \log[P(s)]$ decreases. (b) The normalized avalanche size distribution $P(s)$ narrows as the system starts to recognize the patterns.

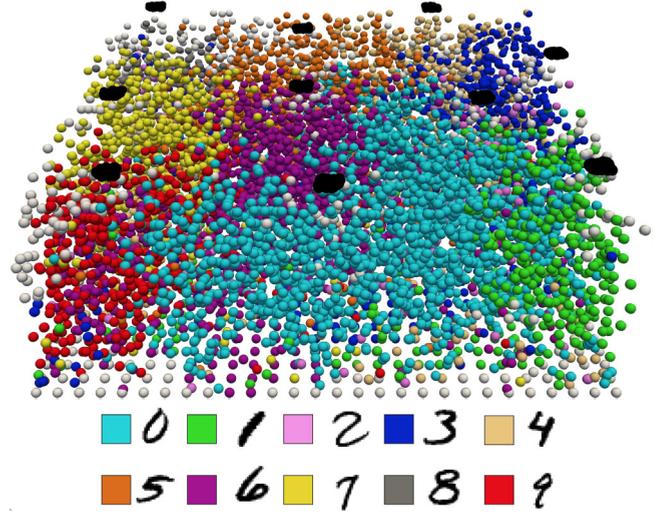


FIG. 5. The network after learning to recognize MNIST patterns. The colors of each neuron indicate for which digit C it has the highest activity $\langle r_i \rangle_C$. Segregated activity regions emerge. The output neurons are shown in black. The parameters used are $N = 8000$, $p_{\text{in}} = 0.3$, $N_O = 50$, $k_{\text{min}} = 10$, $N_C = 10$.

H as the system performs the task. In Fig. 4 we see that as the system learns, the entropy decreases, confirming that the response of the system becomes more and more predictable.

Since in real brains different stimuli lead to different regions being active, we investigate the activity patterns in the network after learning. Let r_i^p be the firing rate of each neuron i for a given input pattern p . We can then define the average firing rate of each neuron $\langle r_i \rangle = 1/I \sum_p r_i^p$ for all input patterns I of all classes in the test set ($I = 10\,000$) and the average firing rate $\langle r_i \rangle_C = 1/I_C \sum_{p \in C} r_i^p$ for each neuron for all the input patterns I_C of a given class C . The $\langle r_i \rangle_C$ are vectors where the i th component represents the firing rate of the i th neuron and their differences can be quantified by calculating the average cosine similarity S_{cos} , defined in Eq. (1) in the Supplemental Material [60]. Figure 3 shows that the cosine similarity of the activity vectors decreases as the performance increases, implying that the firing rates for the various classes are initially very similar and then separate into different activity areas. Figure 1 shows one of the activity regions of a configuration by color coding $\langle r_i \rangle_3 - \langle r_i \rangle$, illustrating the localization of neuronal activity due to class 3. If we ask for which input class each neuron has the highest response, i.e., for which C it has the largest $\langle r_i \rangle_C$, we can assign each neuron to a class. Figure 5 shows that as the system learns to recognize different patterns, segregated activity regions emerge for different pattern classes. Pattern recognition is performed in the visual cortex. Visual stimulus experiments have shown that distinct areas in the primary visual cortex V1 respond to different patterns [51]. For example, a horizontal line stimulus will trigger activity in a different cortical column than a vertical line stimulus [52,53]. Our results show that such activity areas emerge naturally from an initially untrained neural network through the mechanism of negative feedback signals.

Pattern recognition has received wide attention in computer and information science, leading to the development of

efficient algorithms such as the back-propagation method [4]. These algorithms excel at recognition tasks but require each synapse to know its position within the network and the full downstream path to the outputs. This is generally considered biologically implausible [6] and another mechanism based on neurological ingredients is needed to explain pattern recognition in the brain. We have shown that negative feedback signals provide such a mechanism, motivated by the experimental observation that chemical signals such as dopamine [65] mediate synaptic plasticity [39–41]. In this mechanism synapses are only influenced by locally available variables and synaptic adaptation depends solely on the neuronal activity and chemical signals which diffuse through extracellular

space. The presented results focus on the scenario that only the output neurons release chemical error signals. Future research could investigate the addition of dopamine-releasing neurons within the main network, which would allow for a more detailed tuning of the structure and possibly a performance increase. Nevertheless, the network learns to recognize line patterns with high accuracy and performs reasonably well on the MNIST database. The interesting observation of the spatial segregation of activity further supports that learning in biological systems occurs through negative feedback signals.

H.J.H. acknowledges funding from the Brazilian Foundations CAPES and FUNCAP.

-
- [1] S. J. Martin, P. D. Grimwood, and R. G. Morris, *Annu. Rev. Neurosci.* **23**, 649 (2000).
- [2] N. Caporale and Y. Dan, *Annu. Rev. Neurosci.* **31**, 25 (2008).
- [3] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, *Nature (London)* **323**, 533 (1986).
- [4] Y. Bengio, I. J. Goodfellow, and A. Courville, *Nature (London)* **521**, 436 (2015).
- [5] D. Kriesel, A Brief Introduction to Neural Networks Zeta2 edition (2007), available at <http://www.dkriesel.com>.
- [6] F. Crick, *Nature (London)* **337**, 129 (1989).
- [7] S. Grossberg, *Cognit. Sci.* **11**, 23 (1987).
- [8] T. P. Lillicrap, D. Cownden, D. B. Tweed, and C. J. Akerman, *Nat. Commun.* **7**, 13276 (2016).
- [9] Y. Bengio, D.-H. Lee, J. Bornschein, T. Mesnard, and Z. Lin, [arXiv:1502.04156](https://arxiv.org/abs/1502.04156).
- [10] R. D. Brandt and F. Lin, in *Proceedings of the 1996 IEEE International Symposium on Intelligent Control* (IEEE, New York, 1996), pp. 86–90.
- [11] A. Cardoso and A. Wichert, *Neurocomputing* **99**, 575 (2013).
- [12] W. Maass, T. Natschläger, and H. Markram, *Neural Comput.* **14**, 2531 (2002).
- [13] H. Jaeger, German National Research Center for Information Technology, GMD Technical Report No. 148 (2001), p. 13 (unpublished).
- [14] W. Maass, in *Computability in Context: Computation and Logic in the Real World*, edited by S. B. Cooper and A. Sorbi (World Scientific, Singapore, 2011), pp. 275–296.
- [15] D. V. Buonomano and W. Maass, *Nat. Rev. Neurosci.* **10**, 113 (2009).
- [16] D. Plenz and H. G. Schuster, *Criticality in Neural Systems* (Wiley-VCH, Weinheim, 2014).
- [17] D. R. Chialvo, *Nat. Phys.* **6**, 744 (2010).
- [18] J. M. Beggs and D. Plenz, *J. Neurosci.* **23**, 11167 (2003).
- [19] J. M. Beggs and D. Plenz, *J. Neurosci.* **24**, 5216 (2004).
- [20] D. Plenz and T. C. Thiagarajan, *Trends Neurosci.* **30**, 101 (2007).
- [21] O. Shriki, J. Alstott, F. Carver, T. Holroyd, R. N. Henson, M. L. Smith, R. Coppola, E. Bullmore, and D. Plenz, *J. Neurosci.* **33**, 7079 (2013).
- [22] V. Pasquale, P. Massobrio, L. Bologna, M. Chiappalone, and S. Martinoia, *Neuroscience* **153**, 1354 (2008).
- [23] A. Mazzoni, F. D. Broccard, E. Garcia-Perez, P. Bonifazi, M. E. Ruaro, and V. Torre, *PLoS ONE* **2**, e439 (2007).
- [24] T. Petermann, T. C. Thiagarajan, M. A. Lebedev, M. A. Nicolelis, D. R. Chialvo, and D. Plenz, *Proc. Natl. Acad. Sci. USA* **106**, 15921 (2009).
- [25] E. D. Gireesh and D. Plenz, *Proc. Natl. Acad. Sci. USA* **105**, 7576 (2008).
- [26] S. di Santo, P. Villegas, R. Burioni, and M. A. Muñoz, *Proc. Natl. Acad. Sci. USA* **115**, E1356 (2018).
- [27] F. Pittorino, M. Ibáñez-Berganza, M. di Volo, A. Vezzani, and R. Burioni, *Phys. Rev. Lett.* **118**, 098102 (2017).
- [28] P. Bak, C. Tang, and K. Wiesenfeld, *Phys. Rev. Lett.* **59**, 381 (1987).
- [29] L. de Arcangelis, C. Perrone-Capano, and H. J. Herrmann, *Phys. Rev. Lett.* **96**, 028107 (2006).
- [30] L. de Arcangelis, F. Lombardi, and H. Herrmann, *J. Stat. Mech.* (2014) P03026.
- [31] L. Cocchi, L. L. Gollo, A. Zalesky, and M. Breakspear, *Prog. Neurobiol.* **158**, 132 (2017).
- [32] L. de Arcangelis, *Eur. Phys. J.: Spec. Top.* **205**, 243 (2012).
- [33] E. Miranda and H. Herrmann, *Physica A* **175**, 339 (1991).
- [34] L. M. van Kessenich, L. de Arcangelis, and H. Herrmann, *Sci. Rep.* **6**, 32071 (2016).
- [35] L. de Arcangelis and H. J. Herrmann, *Proc. Natl. Acad. Sci. USA* **107**, 3977 (2010).
- [36] V. Capano, H. J. Herrmann, and L. de Arcangelis, *Sci. Rep.* **5**, 9895 (2015).
- [37] D. L. Berger, L. de Arcangelis, and H. J. Herrmann, *Sci. Rep.* **7**, 11016 (2017).
- [38] D. R. Chialvo and P. Bak, *Neuroscience* **90**, 1137 (1999).
- [39] F. Liu, Q. Wan, Z. B. Pristupa, X.-M. Yu, Y. T. Wang, and H. B. Niznik, *Nature (London)* **403**, 274 (2000).
- [40] R. A. Wise, *Nat. Rev. Neurosci.* **5**, 483 (2004).
- [41] J. N. Reynolds and J. R. Wickens, *Neural Networks* **15**, 507 (2002).
- [42] K. Fuxe, A. B. Dahlström, G. Jonsson, D. Marcellino, M. Guescini, M. Dam, P. Manger, and L. Agnati, *Prog. Neurobiol.* **90**, 82 (2010).
- [43] L. Michiels van Kessenich, M. Luković, L. de Arcangelis, and H. J. Herrmann, *Phys. Rev. E* **97**, 032312 (2018).
- [44] M. Watanabe, K. Maemura, K. Kanbara, T. Tamayama, and H. Hayasaki, *Int. Rev. Cytol.* **213**, 1 (2002).
- [45] D. O. Hebb, *The Organization of Behavior: A Neuropsychological Approach* (Wiley, New York, 1949).
- [46] N. S. Simons-Weidenmaier, M. Weber, C. F. Plappert, P. K. Pilz, and S. Schmid, *BMC Neurosci.* **7**, 38 (2006).

- [47] P. S. Kaeser and W. G. Regehr, *Curr. Opin. Neurobiol.* **43**, 63 (2017).
- [48] C. Rosenmund and C. F. Stevens, *Neuron* **16**, 1197 (1996).
- [49] S. O. Rizzoli and W. J. Betz, *Nat. Rev. Neurosci.* **6**, 57 (2005).
- [50] K. Ikeda and J. M. Bekkers, *Proc. Natl. Acad. Sci. USA* **106**, 2945 (2009).
- [51] J. G. Nicholls, A. R. Martin, B. G. Wallace, and P. A. Fuchs, *From Neuron to Brain* (Sinauer Associates, Sunderland, MA, 2001), Vol. 271.
- [52] D. H. Hubel and T. N. Wiesel, *J. Physiol.* **160**, 106 (1962).
- [53] D. H. Hubel, *Eye, Brain, and Vision* (Scientific American Library, Scientific American Books, New York, 1995).
- [54] E. Yacoub, N. Harel, and K. Uğurbil, *Proc. Natl. Acad. Sci. USA* **105**, 10607 (2008).
- [55] T. N. Wiesel and D. H. Hubel, *J. Comput. Neurol.* **158**, 307 (1974).
- [56] V. M. Eguiluz, D. R. Chialvo, G. A. Cecchi, M. Baliki, and A. V. Apkarian, *Phys. Rev. Lett.* **94**, 018102 (2005).
- [57] E. M. Callaway, *Neural Networks* **17**, 625 (2004).
- [58] F. Crick and C. Koch, *Nature (London)* **391**, 245 (1998).
- [59] G. G. Turrigiano and S. B. Nelson, *Nat. Rev. Neurosci.* **5**, 97 (2004).
- [60] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevE.99.010302> for a discussion on how the complexity of the task increases with N_C , for a discussion of the various parameters and their effect on the performance, and for the exact definition of the average cosine similarity S_{\cos} .
- [61] Y. LeCun, <http://yann.lecun.com/exdb/mnist/> (1998).
- [62] D. Cireşan, U. Meier, and J. Schmidhuber, in *Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, New York, 2012), pp. 3642–3649.
- [63] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, *Proc. IEEE* **86**, 2278 (1998).
- [64] P. Y. Simard, D. Steinkraus, and J. C. Platt, in *Proceedings of the Seventh International Conference on Document Analysis and Recognition* (IEEE, New York, 2003), pp. 958–963.
- [65] S. Ikemoto, *Brain Res. Rev.* **56**, 27 (2007).